WHAT IS CLAIMED IS:

1. A method of summarizing a stream of ordered information, comprising:

generating a similarity matrix for the stream of ordered information

decomposing the similarity matrix based on a probabilistic matrix factorization into a plurality of component matrices;

determining, for each component matrix, a representative portion of the stream of ordered information;

extracting the determined representative portions; and

combining the extracted representative portions into a summary of the stream of ordered information.

2. The method of claim 1, wherein the stream of ordered information comprises at least one of at least video information, audio information, still image information, and text information.

3. The method of claim 1, wherein the stream of ordered information comprises a plurality of at least video information, audio information, still image information, and text information

4. The method of claim 1, wherein generating the similarity matrix for the stream of ordered information comprises:

windowing the stream of ordered information;

parameterizing the windowed stream of ordered information; and

determining the similarity matrix of the parameterized windowed stream of ordered information.

5. The method of claim 4, wherein the stream of ordered information comprises at least audio information, and parameterizing the stream of ordered information comprises parameterizing the stream of ordered audio information based on at least one of a STFT Fourier Transform, a Mel-Frequency Cepstral Coefficients Analysis, a spectrogram, a Fast Fourier Transform and wavelet decomposition.

6. The method of claim 4, wherein the stream of ordered information comprises at least video information, and parameterizing the stream of ordered information comprises parameterizing the stream of ordered video information based on at least one of a histogram, ortho-normal projections, deriving a decimated image from DC coefficients of compression macroblocks and discrete cosine transforms.

7. The method of claim 4, wherein the stream of ordered information comprises at least text information, and parameterizing the stream of ordered information comprises parameterizing the stream of ordered text information based on at least one of a sentence, a paragraph, a meta-data information, a term-frequency inverse-document frequency information and part of speech information.

8. The method of claim 1, further comprising determining a number of the component matrices of the ordered information based on a function.

9. The method of claim 8, wherein determining the number of the component matrices of the ordered information based on a function comprises determining the number of the component matrices of the ordered information based on a probabilistic factorization rank of the similarity matrix for the ordered information.

10. The method of claim 1, wherein decomposing the similarity matrix based on a probabilistic matrix factorization into the plurality of component matrices comprises decomposing a similarity matrix S into a plurality of component matrices A_k such that:

$$A_k(i,j) = W(i,k)H(k,j), \text{ and}$$

$$S \cong WH = \sum_{k=1}^{K} A_k ,$$

where:

W(i,k) is an N x K matrix;

H(k,j) is a K x N matrix;

N is a number of separable portions of the stream of ordered information;

K is a number of basis vectors; and

k is an integer.

11. The method of claim 10, wherein determining, for each component matrix, a representative portion of the stream of ordered information comprises determining, for each component matrix, for each of a plurality of candidate excerpts from a subpart of the stream of ordered information that corresponds to that component matrix, a measure of how representative that candidate excerpt is.

12. The method of claim 11, wherein determining the measure of how representative that candidate excerpt is comprises determining an average within-class matrix $\bar{A}_k(q,r)$ for that candidate excerpt.

13. The method of claim 12, wherein determining the average within-class matrix $\bar{A}_k(q,r)$ for that candidate excerpt comprises determining $\bar{A}_k(q,r)$ as:

$$\overline{A}_k(q,r) = \frac{1}{N(r-q)} \sum_{m=q}^{r} \sum_{n=1}^{N} A_k(m,n);$$

where:

N is a number of separable portions of the corresponding subpart of the stream of ordered information;

r is a ending portion of the excerpt; and

q is an starting portion of the excerpt.

14. The method of claim 13, wherein determining, for the plurality of candidate excerpts from the subpart of the stream of ordered information that corresponds to that component matrix, a measure of how representative that candidate excerpt is comprises determining the average within-class component matrix $\bar{A}_k(q,r)$ for excerpts with all possible starting points r.

15. The method of claim 14, wherein determining, for the plurality of candidate excerpts from the subpart of the stream of ordered information that corresponds to that component matrix, a measure of how representative that candidate excerpt is comprises determining the average within-class component matrix $\bar{A}_k(q,r)$ for excerpts with all possible starting points r, each excerpt having a fixed length L between the starting point r and the ending point q.

16. The method of claim 14, wherein determining, for the plurality of candidate excerpts from the subpart of the stream of ordered information that corresponds to that component matrix, a measure of how representative that candidate excerpt is comprises determining the average within-class component matrix $\bar{A}_k(q,r)$ for excerpts with all possible ending points q.

17. The method of claim 14, wherein determining, for the plurality of candidate excerpts from the subpart of the stream of ordered information that corresponds to that component matrix, a measure of how representative that candidate excerpt is comprises determining the average within-class component matrix $\bar{A}_k(q,r)$

for excerpts having a length l_k for that component k, where the length l_k is determined as:

$$l_k = L_T \frac{\sum_{i,j} A_k(i,j)}{\sum_k \sum_{i,j} A_k(i,j)},$$

where L_T is a total length for the combined excerpts.

18. The method of claim 11, wherein determining the measure of how representative that candidate excerpt is comprises determining, for a given component matrix A_k, a score $Q_L^{(k)}(i)$ for the i^{th} starting position of the k^{th} component as:

$$Q_L^{(k)}(i) = \frac{1}{NL} \sum_{m=i}^{i+L} \sum_{n=1}^{N} A_k(m,n),$$

where:

N is a number of separable portions of the corresponding subpart of the stream of ordered information;

i is a starting portion of the excerpt; and

L is a length of the excerpt.

19. The method of claim 18, wherein determining the score $Q_L^{(k)}(i)$ for the i^{th} starting position of the k^{th} component comprises determining a starting point $q_L^{(k)}$ for the excerpt to be extracted from the k^{th} component.

20. The method of claim 19, wherein determining the a starting point $q_L^{(k)}$ for the excerpt to be extracted from the k^{th} component comprises finding the starting point $q_L^{(k)}$ that maximizes the score $Q_L^{(k)}$ for the k^{th} component.

21. The method of claim 1, wherein decomposing the similarity matrix based on a probabilistic matrix factorization into the plurality of component matrices comprises decomposing the similarity matrix into the plurality of component matrices using non-negative matrix factorization.

22. The method of claim 21, wherein decomposing the similarity matrix into the plurality of component matrices using non-negative matrix factorization comprises decomposing a similarity matrix S into a plurality of component matrices A_k using non-negative factorization such that:

$$A_k(i,j) = W(i,k)H(k,j), \text{ and}$$

$$S \cong WH = \sum_{k=1}^{K} A_k ,$$

where:

W(i,k) is an N x K matrix;

H(k,j) is a K x N matrix;

N is a number of separable portions of the stream of ordered information;

K is a number of basis vectors; and

k is an integer.

23. The method of claim 22, wherein generating the similarity matrix for the stream of ordered information comprises generating the similarity matrix S such that:

$$L_{NMF} = \sum_{i,j} S(i,j)\log(S_b(i,j)) - S_b(i,j), \text{ and}$$

$$S \cong S_b = WH;$$

where S_b is a linear approximation of the similarity matrix S.

24. The method of claim 23, wherein generating the similarity matrix for the stream of ordered information further comprises determining a similarity value S(i,j)for each position (i,j) in the similarity matrix S as:

$$S(i,j) = \exp(d_c(v_i, v_j) - 1).$$

where:

v_i and v_j are parameter vectors generated from the stream of ordered information for i^{th} and j^{th} portions of the stream of ordered information; and

d_c is the cosine angle between the parameter vectors v_i and v_j.

25. A storage medium storing a set of program instructions executable on a data processing device and usable to summarize a stream of ordered information, the set of program instructions comprising:

instructions for generating a similarity matrix for the stream of ordered information

instructions for decomposing the similarity matrix based on a probabilistic matrix factorization into a plurality of component matrices;

instructions for determining, for each component matrix, a representative portion of the stream of ordered information;

instructions for extracting the determined representative portions; and

instructions for combining the extracted representative portions into a summary of the stream of ordered information.

26. The storage medium of claim 25, wherein the stream of ordered information comprises at least one of at least video information, audio information, still image information, and text information.

27. The storage medium of claim 25, wherein the stream of ordered information comprises a plurality of at least video information, audio information, still image information, and text information

28. The storage medium of claim 25, wherein generating the similarity matrix for the stream of ordered information comprises:

instructions for windowing the stream of ordered information;

instructions for parameterizing the windowed stream of ordered information; and

instructions for determining the similarity matrix of the parameterized windowed stream of ordered information.

29. The storage medium of claim 28, wherein the stream of ordered information comprises at least audio information, and the instructions for parameterizing the stream of ordered information comprise instructions for parameterizing the stream of ordered audio information based on at least one of a STFT Fourier Transform, a Mel-Frequency Cepstral Coefficients Analysis, a spectrogram, a Fast Fourier Transform and wavelet decomposition.

30. The storage medium of claim 28, wherein the stream of ordered information comprises at least video information, and the instructions for parameterizing the stream of ordered information comprise instructions for parameterizing the stream of ordered video information based on at least one of a histogram, ortho-normal projections, deriving a decimated image from DC coefficients of compression macroblocks and discrete cosine transforms.

31. The storage medium of claim 28, wherein the stream of ordered information comprises at least text information, and the instructions for parameterizing the stream of ordered information comprise instructions for

parameterizing the stream of ordered text information based on at least one of a sentence, a paragraph, a meta-data information, a term-frequency inverse-document frequency information and part of speech information.

32. The storage medium of claim 25, further comprising determining a number of the component matrices of the ordered information based on a function.

33. The storage medium of claim 32, wherein the instructions for determining the number of the component matrices of the ordered information based on a function comprise instructions for determining the number of the component matrices of the ordered information based on a probabilistic factorization rank of the similarity matrix for the ordered information.

34. The storage medium of claim 25, wherein the instructions for decomposing the similarity matrix based on a probabilistic matrix factorization into the plurality of component matrices comprise instructions for decomposing a similarity matrix S into a plurality of component matrices A_k such that:

$$A_k(i,j) = W(i,k)H(k,j), \text{ and}$$

$$S \cong WH = \sum_{k=1}^{K} A_k ,$$

where:

W(i,k) is an N x K matrix;

H(k,j) is a K x N matrix;

N is a number of separable portions of the stream of ordered information;

K is a number of basis vectors; and

k is an integer.

35. The storage medium of claim 34, wherein the instructions for determining, for each component matrix, a representative portion of the stream of ordered information comprise instructions for determining, for each component matrix, for each of a plurality of candidate excerpts from a subpart of the stream of ordered information that corresponds to that component matrix, a measure of how representative that candidate excerpt is.

36. The storage medium of claim 35, wherein the instructions for determining the measure of how representative that candidate excerpt is comprise

instructions for determining an average within-class matrix $\bar{A}_k(q,r)$ for that candidate excerpt.

37. The storage medium of claim 36, wherein the instructions for determining the average within-class matrix $\bar{A}_k(q,r)$ for that candidate excerpt comprise instructions for determining $\bar{A}_k(q,r)$ as:

$$A_k(q,r) = \frac{1}{N(r-q)} \sum_{m=q}^{r} \sum_{n=1}^{N} A_k(m,n);$$

where:

N is a number of separable portions of the corresponding subpart of the stream of ordered information;

r is a ending portion of the excerpt; and

q is an starting portion of the excerpt.

38. The storage medium of claim 37, wherein the instructions for determining, for the plurality of candidate excerpts from the subpart of the stream of ordered information that corresponds to that component matrix, a measure of how representative that candidate excerpt is comprise instructions for determining the average within-class component matrix $\bar{A}_k(q,r)$ for excerpts with all possible starting points r.

39. The storage medium of claim 38, wherein the instructions for determining, for the plurality of candidate excerpts from the subpart of the stream of ordered information that corresponds to that component matrix, a measure of how representative that candidate excerpt is comprise instructions for determining the average within-class component matrix $\bar{A}_k(q,r)$ for excerpts with all possible starting points r, each excerpt having a fixed length L between the starting point r and the ending point q.

40. The storage medium of claim 39, wherein the instructions for determining, for the plurality of candidate excerpts from the subpart of the stream of ordered information that corresponds to that component matrix, a measure of how representative that candidate excerpt is comprise instructions for determining the average within-class component matrix $\bar{A}_k(q,r)$ for excerpts with all possible ending points q.

41. The storage medium of claim 38, wherein the instructions for determining, for the plurality of candidate excerpts from the subpart of the stream of ordered information that corresponds to that component matrix, a measure of how representative that candidate excerpt is comprise instructions for determining the average within-class component matrix $\tilde{A}_k(q,r)$ for excerpts having a length l_k for that component k, where the length l_k is determined as:

$$l_k = L_T \frac{\sum_{i,j} A_k(i,j)}{\sum_k \sum_{i,j} A_k(i,j)},$$

where L_T is a total length for the combined excerpts.

42. The storage medium of claim 35, wherein the instructions for determining the measure of how representative that candidate excerpt is comprise instructions for determining, for a given component matrix A_k, a score $Q_L^{(k)}(i)$ for the i^{th} starting position of the k^{th} component as:

$$Q_L^{(k)}(i) = \frac{1}{NL} \sum_{m=i}^{i+L} \sum_{n=1}^{N} A_k(m,n),$$

where:

N is a number of separable portions of the corresponding subpart of the stream of ordered information;

i is a starting portion of the excerpt; and

L is a length of the excerpt.

43. The storage medium of claim 42, wherein the instructions for determining the score $Q_L^{(k)}(i)$ for the i^{th} starting position of the k^{th} component comprise instructions for determining a starting point $q_L^{(k)}$ for the excerpt to be extracted from the k^{th} component.

44. The storage medium of claim 43, wherein the instructions for determining the a starting point $q_L^{(k)}$ for the excerpt to be extracted from the k^{th} component comprise instructions for finding the starting point $q_L^{(k)}$ that maximizes the score $Q_L^{(k)}$ for the k^{th} component.

45. The storage medium of claim 25, wherein the instructions for decomposing the similarity matrix based on a probabilistic matrix factorization into the plurality of component matrices comprise instructions for decomposing the

similarity matrix into the plurality of component matrices using non-negative matrix factorization.

46. The storage medium of claim 45, wherein the instructions for decomposing the similarity matrix into the plurality of component matrices using non-negative matrix factorization comprise instructions for decomposing a similarity matrix S into a plurality of component matrices A_k using non-negative factorization such that:

$$A_k(i,j) = W(i,k)H(k,j), \text{ and}$$

$$S \cong WH = \sum_{k=1}^{K} A_k ,$$

where:

W(i,k) is an N x K matrix;

H(k,j) is a K x N matrix;

N is a number of separable portions of the stream of ordered information;

K is a number of basis vectors; and

k is an integer.

47. The storage medium of claim 46, wherein the instructions for generating the similarity matrix for the stream of ordered information comprise instructions for generating the similarity matrix S such that:

$$L_{NMF} = \sum_{i,j} S(i,j)\log(S_b(i,j)) - S_b(i,j) , \text{ and}$$

$$S \cong S_b = WH;$$

where S_b is a linear approximation of the similarity matrix S.

48. The storage medium of claim 47, wherein the instructions for generating the similarity matrix for the stream of ordered information further comprise instructions for determining a similarity value S(i,j)for each position (i,j) in the similarity matrix S as:

$$S(i,j) = \exp(d_c(v_i, v_j) - 1 .$$

where:

v_i and v_j are parameter vectors generated from the stream of ordered information for i^{th} and j^{th} portions of the stream of ordered information; and

d_c is the cosine angle between the parameter vectors v_i and v_j.

49. A stream of ordered information summarizing system, comprising

a similarity matrix determining circuit, routine or application that determines a similarity between two portions of the stream of ordered information and that stores the determined similarity into a similarity matrix;

a probabilistic decomposition circuit, routine or application that decomposes the similarity matrix into a plurality of component matrices based on a probabilistic matrix factorization;

an excerpt determining circuit, routine or application that determines, for each component matrix, a representative portion of the stream of ordered information;

an excerpt extracting circuit, routine or application that extracts the determined representative portions; and

a summary generating circuit, routine or application that combines the extracted representative portions into a summary of the stream of ordered information.

50. The stream of ordered information summarizing system of claim 49, further comprising:

a windowing circuit, routine or application that windows the stream of ordered information; and

a parameterization circuit, routine or application that parameterizes the windowed stream of ordered information;

wherein the similarity matrix determining circuit, routine or application that determines the similarity between two portions of the windowed and parameterized stream of ordered information.